

What is Claimed is:

1 ~~1. A method for assignment of base numbers to peaks within an experimental~~
2 ~~DNA sequencing data trace derived from the separation of experimental DNA sequencing~~
3 ~~fragments, comprising the steps of:~~

4 ~~(a) obtaining one or more reference DNA sequencing data traces derived from~~
5 ~~the separation of reference DNA sequencing fragments reflecting the position of at least one base~~
6 ~~in a reference polynucleotide of known sequence;~~

7 ~~(b) evaluating the reference DNA sequencing data traces to determine a~~
8 ~~corrected time scale indicative of migration times at which peaks should occur;~~

9 ~~(c) sampling the experimental DNA sequencing data trace at time points~~
10 ~~determined by the corrected time scale, and~~

11 ~~(d) assigning a base number to each peak found in the experimental DNA~~
~~sequencing data trace based upon the corrected time scale.~~

2. The method of claim 1, wherein the step of evaluating the reference DNA
sequence data traces includes the steps of:

(i) identifying a plurality of peaks in the reference DNA sequencing data
traces, and creating a data table containing the number of each peak based on the known
sequence of the polynucleotide, and the position of each peak in the reference DNA sequencing
data trace;

(ii) identifying a set of coefficients for a polynomial effective to substantially
linearize a plot of peak number versus separation between adjacent peaks; and

(iii) creating from the coefficients and the polynomial a corrected time scale
which reflects the positions at which a peak should occur at any given point in a sequencing data
trace.

1 ~~SUB A4~~ 3. The method of claim 1, wherein the experimental DNA sequencing data
2 trace and a first reference DNA sequencing data trace are derived from analysis of sequencing
3 fragments in a common lane of a sequencing gel.

1 4. The method of claim 1, wherein a plurality of reference DNA sequencing
2 data traces are obtained, each derived from the separation of the same set of reference DNA
3 sequencing fragments.

1 ~~SUB A5~~ 5. The method of claim 1, wherein the polynomial is a third or higher order
2 polynomial.

6. The method of claim 1, wherein a defined number of bands are selected
for evaluation from each of the reference DNA sequencing data traces.

7. The method of claim 6, wherein the defined number of bands selected is
from 3 to 40.

8. The method of claim 6, wherein the defined number of bands is at least
equal to the order of the polynomial, plus 1.

1 9. The method of claim 1, wherein base numbers are assigned to peaks
2 within a plurality of experimental DNA sequencing data traces derived from the separation of
3 experimental DNA sequencing fragments indicative of the positions of a plurality of types of
4 bases.

1 10. The method of claim 9, wherein base numbers are assigned to peaks
2 within four experimental DNA sequencing data traces derived from the separation of
3 experimental DNA sequencing fragments indicative of the positions of four types of bases.

1 11. A method for evaluating the sequence of a target polynucleotide,
2 comprising the steps of:

3 (a) obtaining one or more experimental DNA sequencing data traces derived
4 from the separation of experimental DNA sequencing fragments reflecting the position of at least
5 one base in the target polynucleotide and one or more reference DNA sequencing data traces
6 derived from the separation of reference DNA sequencing fragments reflecting the position of at
7 least one base in a reference polynucleotide of known sequence;

8 (b) evaluating the reference DNA sequencing data traces to determine a
9 corrected time scale indicative of migration times at which peaks should occur;

10 (c) sampling the experimental DNA sequencing data traces at time points
11 determined by the corrected time scale, and

12 (d) assigning a base number to each peak found in the experimental DNA
13 sequencing data traces based upon the corrected time scale, thereby obtaining information about
14 the sequence of the target polynucleotide.

1 12. The method of claim 11, wherein the step of evaluating the reference DNA
2 sequence data traces includes the steps of:

3 (i) identifying a plurality of peaks in the reference DNA sequencing data
4 traces, and creating a data table containing the number of each peak based on the known
5 sequence of the polynucleotide, and the position of each peak in the reference DNA sequencing
6 data trace;

7 (ii) identifying a set of coefficients for a polynomial effective to substantially
8 linearize a plot of peak number versus separation between adjacent peaks; and

9 (iii) creating from the coefficients and the polynomial a corrected time scale
10 which reflects the positions at which a peak should occur at any given point in a sequencing data
11 trace.

1 13. ~~The method of claim 11, wherein the reference DNA sequencing traces~~
2 ~~and the experimental DNA sequencing data trace are derived from analysis of sequencing~~
3 ~~fragments in a common sequencing gel.~~

1 14. ~~The method of claim 13, wherein the experimental DNA sequencing data~~
2 ~~trace and a first reference DNA sequencing data trace are derived from analysis of sequencing~~
3 ~~fragments in a common lane of the common sequencing gel.~~

1 15. The method of claim 11, wherein a plurality of reference DNA sequencing
2 data traces are obtained, each derived from the separation of the same set of reference DNA
3 sequencing fragments.

1 16. ~~The method of claim 11, wherein the polynomial is a third or higher order~~
2 ~~polynomial.~~

1 17. ~~The method of claim 11, wherein a defined number of bands are selected~~
2 ~~for evaluation from each of the reference DNA sequencing data traces.~~

1 18. The method of claim 17, wherein the defined number of bands selected is
2 from 3 to 40.

1 19. The method of claim 17, wherein the defined number of bands is at least
2 equal to the order of the polynomial, plus 1.

1 20. The method of claim 11, wherein base numbers are assigned to peaks
2 within a plurality of experimental DNA sequencing data traces derived from the separation of
3 experimental DNA sequencing fragments indicative of the positions of a plurality of types of
4 bases.

1 21. An apparatus for evaluating the sequence of a target polynucleotide,
2 comprising:

3 (a) an input for receiving information about one or more experimental DNA
4 sequencing data traces derived from the separation of experimental DNA sequencing fragments
5 reflecting the position of at least one base in the target polynucleotide and one or more reference
6 DNA sequencing data traces derived from the separation of reference DNA sequencing
7 fragments reflecting the position of at least one base in a reference polynucleotide of known
8 sequence;

9 (b) a processor, operatively programmed to evaluate the reference DNA
10 sequencing data traces to determine a corrected time scale indicative of migration times at which
11 peaks should occur;

12 (c) a processor, operatively programed to sample the experimental DNA
13 sequencing data traces at time points determined by the corrected time scale;

14 (d) a processor, operatively programmed to assign a base number to each peak
15 found in the experimental DNA sequencing data traces based upon the corrected time scale,
16 thereby obtaining information about the sequence of the target polynucleotide; and

17 (e) an output for communicating the information about the sequence of the
18 target polynucleotide.

1 22. The apparatus of claim 21, wherein the processor programmed to evaluate
2 the reference DNA sequence data traces is programmed to perform the steps of:

3 (i) identifying a plurality of peaks in the reference DNA sequencing data
4 traces, and creating a data table containing the number of each peak based on the known
5 sequence of the polynucleotide, and the position of each peak in the reference DNA sequencing
6 data trace;

7 (ii) identifying a set of coefficients for a polynomial effective to substantially
8 linearize a plot of peak number versus separation between adjacent peaks; and

9 (iii) creating from the coefficients and the polynomial a corrected time scale
10 which reflects the positions at which a peak should occur at any given point in a sequencing data
11 trace.

00420-FF5E50